



University of Groningen

Target selection for direct marketing.

Bult, Jan Roelf

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1993

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bult, J. R. (1993). Target selection for direct marketing. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 7

Summary, managerial implications, and future research

In this thesis we concentrated on the use of direct mail for targeting potential buyers. The major characteristics that influences the success of a promotional direct mail campaign are the offer, the communication elements, the timing or sequence of these communication elements, and the list of customers to be targeted. Of these characteristics, the list is considered to be the most important one in order to stimulate the response to an offer. When the message does not reach the proper targets, it has little chance of being effective. Therefore, we did consider the process of selecting those individuals whose probability of response to the mailing is sufficiently high.

7.1 Summary

In this thesis we concentrated on the use of direct mail for targeting consumers, although the methodology presented can also be used for other types of media and targeting industrial buyers. The objective of this thesis is to:

1. evaluate the methods discussed in the direct marketing literature and used by direct marketing firms for the selection of targets,
2. propose such modifications in methods that they can be used more effectively for target selection for direct marketing, and
3. develop new techniques which eliminate certain disadvantages of the methods under (1) and (2).

The lists, variables, and techniques used by direct marketing firms for the selection of targets have been discussed in chapter 2. In order to

select targets, suppliers using direct mail have to make several decisions. First, they have to choose among available lists from which individuals can be targeted. They might, for example, use their internal list (houselist) or they might also rent external lists; some of those are primarily compiled for direct marketing purposes. Examples of these lists are postcode information systems. Next, after one or more lists have been selected, they have to select selection variables discriminating responders from non-responders. The (mailing) lists and selection variables most frequently used for the selection of targets were discussed in chapter 2. The internal list (if available) is considered to be the most important list. This has to do with the general direct marketing knowledge that the 'best' predictors for future behavior are variables measuring behavior in the past. Therefore, it is not surprising either that, in general, the most important selection variables are Recency of last purchase, Frequency of purchase, and Monetary value spent on the purchases (RFM) variables. When information is wanted on all households in the Netherlands, geographically based selection variables can be used. These variables are aggregated at the geographic level of the postcode. Finally, suppliers have to choose between available statistical selection techniques. The techniques most frequently used for identifying the best targets are (see also chapter 2): cross-tabulation, (CH)AID, discriminant analysis, the linear probability model, the logit model, and the probit model. In some studies, individuals are ranked (from the highest to the lowest score) according to their estimated probability of response. Next, groups (mostly deciles) of equal size are defined for which the actual average probability of response is computed. By means of such a decile report, the supplier decides which deciles should be mailed and which deciles should not. This selection procedure is known as *gains chart analysis*.

In chapter 3 we proposed modifications in the gains chart analysis. Since there appears to be not a single paper discussing the analytical and statistical aspects involved in the gains chart analysis, we presented in chapter 3 a comprehensive framework for the selection of targets for a promotional direct mail campaign. Furthermore, most of the traditional selection methods do not consider the question of an optimal selection strategy. By equating marginal returns and marginal costs we introduced a method to determine which households should be mailed. Two important curves were derived. Firstly, we developed the response curve, which describes the relationship between the fraction of 'best' households being selected and the average probability of response in this selected group. Secondly, the cut-off curve was developed. This curve describes the relationship between

the ratio of returns per successful reply and the costs per mailing, and the optimal fraction of people that should be mailed.

The functional forms of both these curves heavily rest upon the functional form of the distribution function of the disturbance terms. Therefore, we employed a semiparametric estimation method. The framework introduced in chapter 3 added three aspects to the gains chart analysis used in the direct marketing literature so far.

- Firstly, we explicitly take the maximization principle for the selection of targets. This means that on average our framework is at least as good as the existing frameworks. This means that the analysis presented in chapter 3 yields, on average, higher monetary returns than existing ones.
- Secondly, our approach offers an integrated theoretical frame that can be put to work in other contexts as well. In chapter 4 we presented, by means of examples, three extensions of our framework.
- Thirdly, the response curve we derived is by construction non-increasing over the whole range of possible values of the fraction of households being selected. As a result, the cutoff curve is non-decreasing. Hence a particular value of the ratio of returns per positive reply and the costs per mailing is guaranteed to give a unique profit maximizing value for the optimal fraction of people that should be mailed.

In chapter 4, we presented three extensions of gains chart analysis and developed new techniques for situations which are beyond the scope of the analysis presented in chapter 3.

- Firstly, we considered the situation in which there are measurement errors in the explanatory variables. When using, for example, data from postcode information systems, measurement errors are introduced due to the aggregation level. We discussed the possible effects of measurement errors by studying a very simple example.
- Secondly, elaborating on an example given in chapter 3 we discussed the situation in which different mailings (of different qualities) with different costs per mailing are sent, offering the same product or service at the same time to different people. A multilayer response curve was derived. Whether a multilayer mailing strategy is optimal depends on the trade-off between higher costs per mailing (for high quality mailings) and the increase in the average probability of response towards the mailing.

- Thirdly, we discussed situations in which we are not so much interested in modeling response probabilities but in the purchase amount involved in positive replies.

In chapter 5, we studied various parametric and semiparametric models in order to obtain an empirical specification for the response curve and the cutoff curve. The benefit of semiparametric models in general is that only weak distributional assumptions are made. Therefore, these models are more robust against specification error and the resulting parameter estimates are consistent under weak assumptions. Furthermore, more allowance is made for the information contained in the data. These models can, for example, allow for heteroskedasticity of the disturbance terms of unknown functional form. However, there are several limitations associated with semiparametric models. The objective function is usually not smooth and there is in general no closed form solution for the estimators of the parameters. Since less structure is added to the relation between the variables, the variance (root mean square error) of the estimated parameters increases. Also testing hypotheses is in general not straightforward in semiparametric models.

We distinguished two types of parametric models. Firstly, we studied parametric models that assume a homogeneous relationship between the dependent variable and the independent variables. The parametric models most frequently used are the linear probability model, the logit model, and the probit model. Secondly, we discussed a latent class logit model in which it is assumed that the market consists of an unknown number of latent classes each having their own relationship between the dependent variable and the independent variables. In our empirical example of the latent class logit model we found two segments with very different coefficients. However, most of the households had an almost equal probability of belonging to segment one as they had for segment two.

While in chapter 3 we only estimated the distribution function of the disturbance terms semiparametrically, we estimated in this chapter also the unknown parameters according to the semiparametric estimation method introduced by Cosslett (1983).

In chapter 6, we looked at a different approach for the selection of targets. Instead of using gains chart analysis, we directly classified each household into one of the two categories (mail-receiver or non-receiver). We presented methods, which were based on minimizing loss functions. Three different loss function were considered: (1) the symmetric loss function, (2) the asymmetric homogeneous loss function, and (3) the asymmetric heterogeneous loss function. Symmetric loss functions assume that the costs

of misclassification are the same for responders and non-responders. Asymmetric loss function assume that the costs of misclassification are not equal. A direct marketing company decides for every available address whether it should receive an offer by mail or not. Misclassification of a household that would have responded to the mailing but did not receive the mailing carries a greater loss than misclassification of a household that received the mailing but did not respond. In case of an asymmetric heterogeneous loss function, it is assumed that the costs of misclassification are asymmetric and household-specific. Parametric as well as semiparametric classification models were employed. The semiparametric models used were the maximum score method of Manski (1975) and the smoothed maximum score function of Horowitz (1992). These semiparametric have not yet been applied for the selection of targets. The only assumption needed to identify the parameters in these models is that one quantile of the dependent variable y_i is known to be linear in x_i (the vector of independent variables). We showed that this assumption is very closely related to the asymmetric loss function. In order to compare the maximum score method with the linear probability model, the logit model, and the probit model, we tested whether the estimated parameters were significantly different from each other or not. We also computed the predictive accuracy of the models. It turns out that the estimates of the parametric models differ significantly from the estimates of maximum score. In case of the probit model, the normality assumption was also rejected. Based on the findings of Manski and Thompson (1986), the difference between the estimates could be due to inconsistency of the estimators of the parametric models.

Although the predictive accuracy in the validation sample of maximum score is worst in terms of percentage of correctly classified observations, the maximum score method is best in terms of the asymmetric loss function. When asymmetry is accommodated, the overall classification accuracy necessarily decreases, because the costs of misclassification are heterogeneous. Asymmetric loss functions favor households for whom the costs of misclassification is the highest. It is in this sense that maximum score gives more weight to the most important observations.

We also estimated a second version of maximum score in which the assumption made about the disturbance term depends on the p -value in the asymmetric absolute loss function. The estimated parameters turn out to be enormously different from the other models. In terms of profit maximization, this second version of maximum score performs much better than the models most frequently used. From this it is clear that market

researchers who face an asymmetric absolute loss function should use the maximum score method instead of the parametric models.

In section 6.4, we discussed the asymmetric heterogeneous loss function and its relationship to the smoothed maximum score estimator. We compared this estimator empirically with its homogeneous counterpart, and we saw that profit is lost when using an asymmetric homogeneous loss function in case the costs of misclassification are heterogeneous distributed.

7.2 Managerial implications

In this section, we discuss implications for direct marketing management which are concerned with the selection of targets. Which methodology the firm should use depends on the type of mailing. The different types of mailing discussed are: (1) a single mailing with just one offer, (2) multiple mailings with just one offer, and (3) mailings (single or multiple) where the returns per positive reply are household-specific.

In case of a single mailing with just one offer, we have a situation in which the ratio of costs and returns per positive reply are constant. The marketing researcher can use either the gains-chart methodology or minimizing the homogeneous asymmetric loss function. In case of gains-chart analysis, the marketing researcher is able to visualize the selection process by the response curve and the cutoff curve. These curves depend on the functional form of the disturbance terms. This form can either be estimated by parametric or semiparametric estimation methods. Parametric methods are easier to use (and understand) than semiparametric methods, but the use of parametric methods might lead to inconsistent estimators of the functional form of the disturbance term. Therefore, we advocate the use of semiparametric estimation methods. In case of minimizing loss functions, semiparametric classification methods are preferred in terms of profit maximization.

If the firm uses more than one mailing offering the same (single) product, it should use the multilayer strategy discussed in section 4.2. The methodology presented in this section also determines whether using multiple mailings are preferred over using a single mailing.

If the ratio of costs per mailing and returns per positive reply is household-specific, the firm could either use the gains-chart methodology briefly discussed in section 4.3 or use the strategy of minimizing a heterogeneous loss function. This last strategy is only useful if the marketing researcher is able to specify in advance which household belong to which